DOCUMENT RESUME

ED 151 381 TM 006 799

AUTHOR Cahen, Leonard S.
TITLE Report on "A" Testing. Beginning Teacher Evaluation Study. Technical Note Series. Technical Note III-2.
INSTITUTION Far West Lab. for Educational Research and Development, San Francisco, Calif.
PUB DATE 22 Dec 76
NOTE 11p.; For related documents, see TM 006 800 and 801

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Academic Achievement; *Achievement Tests; Affective Tests; Complexity Level; Elementary Education; *Elementary School Mathematics; Grade 2; Grade 5; Item Analysis; *Reading; *Statistical Data; Student Testing; Teacher Evaluation; *Testing; *Testing Programs; Test Interpretation

ABSTRACT

The "A" testing program of the Beginning Teacher Evaluation Study, Phase III-B, is summarized. A sample of 150 second grade and 134 fifth grade students completed the reading and mathematics achievement battery, as well as affective scales regarding reading, mathematics, and school. The administration procedures, test results, and psychometric properties of the tests are described. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED151381

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

BTES

BEGINNING TEACHER EVALUATION STUDY

Technical Note Series

Technical Note III-2

REPORT ON "A" TESTING

December 23, 1976

FAR WEST LABORATORY
FOR EDUCATIONAL RESEARCH AND DEVELOPMENT
1855 FOLSOM STREET • SAN FRANCISCO, CALIFORNIA 94103 • (415) 565-3000

TM 006799

Technical Note III-2

Report on "A" Testing

by

Leonard S. Cahen

December 22, 1976

Overview

This report describes the "A" testing which took place between October 18, 1976 and November 5, 1976. This period included the regularly scheduled testing dates plus make-up testing for "target students" who were absent for the scheduled test administration.

"Target students" from 25 second grade and 23 fifth grade classrooms participated in the "A" testing. These 48 classrooms were located in ten different school districts. One hundred and thirty-four target students were tested at the fifth grade. There was no attrition of fifth grade students between "Selection" testing and "A" testing. All students selected were tested in this grade. However, during the period between "Selection" and "A" testing, ten second grade "target students" were lost. We replaced these students with students who were designated as "alternates" at the time the original selection of "target students" was undertaken. Thus, we tested 150 second grade "target students" at the time of the "A" testing.

Test Administration

The test administrators had been given one half-day of training to familiarize themselves with the test booklets, directions, time limits, and to obtain an orientation for administering tests to a small number of pupils. Special rooms were set up in each school for the administration of the test. "Target students" were removed from their regular classrooms and tested as a group. After testing, the booklets were returned to the Laboratory for processing. Any rescheduling necessary to test the absent "target students" was undertaken immediately. In general, the test administrators reported that they had better control of the testing situation when testing six students, as opposed to the whole class, as was the case at the administration of the Selection Battery.

Information to Teachers

After data were run through the item analysis program which provided a check on the scoring procedures, keys, etc. psychometric properties, including the total score distributions were examined for each of the sub-scales. A format was developed to provide a computerized printout of scores for each teacher listing information for each of their "target students." A Manual of Interpretation was also developed. Each teacher was mailed a copy of scores for her students plus the Manual of Interpretation. These materials were sent to the teachers by Certified Mail (return receipt requested). We were able to maintain our policy of confidentiality of data by having the test scores returned directly to the appropriate teacher. We were also able to keep our promise to teachers that we would share our data with them.

Psychometric Properties of the Tests

Tables 1 and 2 display summary psychometric data for the grade 2 Reading and Mathematics subtests. Tables 3 and 4 display summary psychometric data for the fifth grade Reading and Mathematics subtests. All scores are corrected for chance guessing.

Insert Tables 1-4 Here

In Tables 1 through 4 the symbol $\overline{P}$ serves as an index of the overall difficulty of the test. This statistic is defined as the subtest mean divided by the number of items. In some cases the reliability coefficient (Alpha) may appear to be low. The reader is reminded that the procedure used for the selection of "target students" reduced the variability in ability significantly for these students (as compared to the total population of students we tested at grades 2 and 5 at the time of Selection). The reduction in the variability has probably led to a reduced magnitude of alpha. The standard error of

measurement is also presented in these tables. The standard error of measurement is a statistic that should not be dramatically reduced when the variability of students is reduced. A key statistic shown in Tables 1 through 4 is the percent of students passing 80 percent or more of the items in the subtest or scale. A major goal in the Selection of students to observe was to reduce the number of students who would obtain perfect or near perfect on our scales as they progressed through Phase III-B. The Selection testing appears to have performed that function. This statistic will be watched carefully in future testings. The speed test in multiplication, which is shown in Table 4, had 62 percent of the students passing 80 percent or more of the items in this scale at the "A" testing. We immediately tripled the number of items for this scale so that ceiling effects will be negligible in future testings.

Table 5 shows the summary psychometrics of the three Affective scales at grade 2 and 5. At second grade, each item can earn a score of 1 to 3 points. At grade 5, the "Likert" type items have five response categories. This would mean that the possible range of scores on the reading affect scale at grade 2 would be from 7 to 21. The reading scale at grade 5 would have a possible range of 7 to 35. In general, the staff feels that the testing process is proceeding smoothly.

Insert Table 5 Here

Table 1

Psychometric Data
Grade 2 Reading Subtests
"A" Administration
(N = 150)

Subtests	No. of Items	$\bar{X}$	σ	$\bar{P}$	Alpha	Standard Error of Measurement	Percent Passing 80% or more of Items in Subtest
Decoding Consonant Blends	12	6.79	3.63	.57	.69	1.94	23
Decoding Variant Consonants	14	4.41	5.60	.32	.72	2.83	15
Decoding Long Vowels	20	7.89	6.05	.40	.78	2.79	10
Decoding Complex Patterns	16	4.86	4.28	.30	.69	2.30	1
Compound Words	8	4.71	2.26	.59	.77	1.01	25
Root Words and Affixes - D	9	3.87	1.81	.43	.53	1.17	0
Root Words and Affixes - C	8	1.59	2.83	.20	.60	1.68	1
Syllables	10	2.43	2.68	.24	.36	2.04	0
Synonyms	16	4.36	4.48	.27	.77	2.06	1
Pronoun Reference	10	1.91	2.74	.19	.41	1.99	0

Table 2

Psychometric Data
Grade 2 Mathematics Subtests
"A" Administration
(N = 150)

Subtests	No. of Items	$\bar{X}$	σ	$\bar{P}$	Alpha	Standard Error of Measurement	Percent Passing 80% or more of Items in Subtest
Money	10	4.24	2.91	.42	.62	1.71	13
Linear Measurement	11	6.84	2.05	.62	.36	1.57	15
Addition No Regrouping	16	9.08	3.51	.57	.82	1.84	25
Speed Test - Addition	14	5.07	3.14	.36	.74	1.55	0
Place Value/ Numerals	25	5.50	4.63	.22	.66	2.66	0
Subtraction No Regrouping	16	6.01	4.39	.38	.76	2.09	1
Word Problems	15	3.05	2.69	.20	.45	2.13	0
Computational Transfer	22	4.05	4.62	.18	.64	2.70	0
Speed Test - Subtraction	14	4.08	3.15	.30	.66	1.76	0
Addition No Regrouping	4	2.31	1.26	.58	.41	.84	25
Subtraction No Regrouping	4	2.52	1.06	.63	.29	.78	22

Table 3

Psychometric Data
Grade 5 Reading Subtests
"A" Administration
(N = 134)

Subtests	No. of Items	$\bar{X}$	σ	$\bar{P}$	Alpha	Standard Error of Measurement	Percent Passing 80% or more of Items in Subtest
Syllables	21	8.86	6.14	.42	.77	2.84	0
Root Words and Affixes	15	4.65	3.23	.31	.49	2.24	0
Synonyms	15	5.80	4.00	.39	.69	2.13	1
Pronouns	16	7.64	4.00	.14	.67	2.24	13
Word Meaning in Context	20	6.94	5.85	.35	.80	2.57	10
Comprehension - Literal	14	5.86	4.26	.42	.73	2.10	14
Comprehension - Inferential	9	3.90	2.56	.43	.51	1.69	12
Comprehension - Main Idea - Literal	5	2.22	1.79	.44	.45	1.19	14
Comprehension - Main Idea - Inference	4	1.05	1.46	.26	.34	1.03	1
Topic Sentence	10	3.66	3.08	.37	.62	1.81	1
Fact and Opinion	10	1.86	2.71	.19	.50	1.82	0

Table 4

Psychometric Data
Grade 5 Mathematics Subtests
"A" Administration
(N = 134)

Subtests	No. of Items	$\bar{X}$	σ	$\bar{P}$	Alpha	Standard Error of Measurement	Percent Passing 80% or more of Items in Subtest
Addition and Subtraction	10	7.50	2.66	.75	.73	1.31	57
Open Multiplication	4	1.68	1.63	.42	.85	.54	20
Properties and Factors	18	12.03	3.47	.67	.81	1.47	24
Number Patterns	8	2.14	2.20	.27	.44	1.53	0
Multiplication	16	7.55	5.46	.47	.87	1.90	24
Speed Test - Multiplication	14	10.66	4.05	.76	.89	1.31	62
Place Value/ Numerals	10	4.55	2.46	.46	.52	1.62	10
Word Problems	12	2.00	2.91	.17	.53	1.92	0
Perimeters	5	1.22	1.57	.24	.37	1.11	0
Area	5	.26	1.16	.05	.15	.96	0
Lines and Figures	6	.51	1.59	.09	.29	1.22	0
Number Pairs	6	.98	1.87	.16	.47	1.25	0

Table 5

Psychometric Data
Affect Measures
"A" Administration

Grade 2 (N = 150)

Affect Scale	Items	$\bar{X}$	σ	Alpha	Standard Error of Measurement
Reading	7	15.42	3.57	.71	1.77
Mathematics	7	16.04	4.11	.83	1.56
School	2	4.83	1.37	.58	.63

Grade 5 (N = 134)

Affect Scale	Items	$\bar{X}$	σ	Alpha	Standard Error of Measurement
Reading	11	42.11	7.47	.83	2.97
Mathematics	11	38.70	9.88	.90	2.99
School	2	7.52	2.15	.55	1.02